Linear Transformations That Preserve Majorization, Schur Concavity, and Exchangeability

A. M. Dean and J. S. Verducci
Department of Statistics
The Ohio State University
Columbus, Ohio 43210

Submitted by M. D. Perlman

ABSTRACT

A characterization is given for the linear transformations that preserve majorization, Schur concavity, and exchangeability. In the case of bijectio⋅:s, it is shown that the same class of linear transformations preserves all three properties. More generally it is shown that the class of injections preserving majorization consists of the symmetric generalized inverses of the class of surjections that preserve exchangeability. Explicit matrix forms are given for each class, and the theory is applied to determine the partial exchangeability of estimators of interactions as well as to develop two extensions of the Marshall-Olkin theorem.

1. INTRODUCTION

The preordering of majorization, defined formally below, provides a unified framework for the study of inequalities involving functions of several variables. Marshall and Olkin (1979) offer a comprehensive introduction to this topic. Central to the theory of majorization is the work of Schur (1928), who characterized the set of real-valued functions f that reverse the preordering of majorization in the sense that $\mathbf{x} < \mathbf{y}$ implies $f(\mathbf{x}) \geqslant f(\mathbf{y})$. This property, called Schur concavity, occurs in many common multivariate density functions. In the present paper, we extend some of Schur's results to vector-valued functions. In particular, we characterize all linear transformations $A: \Re^m \to \Re^n$ $(m \leqslant n)$ that preserve majorization, giving the explicit matrix form of A. In addition, we detail some implications of these transformations for the joint distributional properties of contrast estimators as well as for multivariate distributions with singular density functions.

655 Avenue of the Americas, New York, NY 10010 0024-3795/90/$3.50

A necessary and sufficient condition for any linear transformation to preserve exchangeability is presented in Condition 2 below. From this condition, the explicit form of linear transformations preserving exchangeability is derived. The relationship among linear transformations that preserve majorization, Schur concavity, and exchangeability is established through a series of theorems. In the case $m = n$, it is shown that the same set of nonsingular linear transformations preserve all three properties.

The basic results for surjections are used to extend further a theorem of Anderson (1955), extended by Marshall and Olkin (1974) and Kimura and Kakiuchi (1989) among others, concerning the decreasing probability content of a region symmetric about the origin as it undergoes a location shift.

In summary, our work characterizes the linear injections that preserve majorization, the linear surjections that preserve exchangeability, and the linear bijections that preserve Schur concavity. The formal proofs for these results are given in Section 2 in the case where the random variables have proper densities. Section 3 contains results about partial exchangeability, including an application concerning the joint distribution of estimators of interactions. The case of singular densities, which have applications to the distribution of robust test statistics, is treated in Section 4.

DEFINITION. A vector $\mathbf{w} \in \Re^n$ is said to *majorize* another vector $\mathbf{x} \in \Re^n$ if

$$\sum_{i=1}^{n} \mathbf{w}_i = \sum_{i=1}^{n} \mathbf{x}_i$$

and

$$\sum_{i=1}^{k} \mathbf{w}_{[i]} \geqslant \sum_{i=1}^{k} \mathbf{x}_{[i]}$$

for each $k = 1, \dots, n-1$, where $\mathbf{x}_{[1]} \geqslant \cdots \geqslant \mathbf{x}_{[n]}$ denote the components of $\mathbf{x}$ in decreasing order. We express such a majorization as $\mathbf{w} > \mathbf{x}$. A subset S of $\Re^n$ is said to be *S-convex* if $\mathbf{x} \in S$ whenever $\mathbf{w} \in S$ and $\mathbf{x} < \mathbf{w}$. A function $f: \Re^n \to \Re$ is *Schur-concave* if and only if $f(\mathbf{w}) \leqslant f(\mathbf{x})$ whenever $\mathbf{w} > \mathbf{x}$.

A *permutation* τ is a one-to-one function from the set $\{1, \dots, n\}$ onto itself. The set of all such permutations is denoted by Ω_n. A *permutation matrix* $R(\tau)$ is an $n \times n$ matrix with the (i, j)th entry equal to 1 if $\tau(i) = j$ and 0 otherwise. A random vector $\mathbf{x} \in \Re^n$ is said to have *exchangeable* *components* if $R(\tau)\mathbf{x}$ has the same distribution as $\mathbf{x}$ for every permutation $\tau \in \Omega_n$. A linear transformation $A: \Re^m \to \Re^n$ *preserves majorization* if $A\mathbf{y} > A\mathbf{z}$ whenever $\mathbf{y} > \mathbf{z}$. A linear transformation B of a random vector $\mathbf{x}$ is said to *preserve Schur concavity* if the density of $B\mathbf{x}$ is Schur-concave whenever the density of $\mathbf{x}$ itself is Schur-concave; similarly B is said to *preserve exchangeability* if $B\mathbf{x}$ has exchangeable components whenever $\mathbf{x}$ does.

We also adopt the following notational conventions. The vector $\mathbf{e}_n \in \Re^n$ is the vector of ones; that is, $\mathbf{e}_n = (1, \dots, 1)'$. The matrix I_n is the $n \times n$ identity matrix, and $J_n = \mathbf{e}_n \mathbf{e}_n'$. We shall also always assume that the positive integers m and n satisfy $m \leqslant n$, and that linear transformations $A: \Re^m \to \Re^n$ and $B: \Re^n \to \Re^m$ are of full rank m. Under these assumptions, A and B are *generalized inverses* if $BA = I_m$ and $AB = (AB)'$. A generalized inverse $B = A^-$ of A always exists and, under the conditions of full rank and symmetry stated above, is unique (see e.g. Graybill, 1983, Chapter 6).

For general full-rank linear transformations $A: \Re^m \to \Re^n$ and $B: \Re^n \to \Re^m$, the following two conditions play an important role in the sequel.

Condition 1. For every permutation $\pi \in \Omega_m$ there exists a permutation $\tau \in \Omega_n$ such that $R(\tau)A = AR(\pi)$.

Condition 2. For every permutation $\pi \in \Omega_m$ there exists a permutation $\tau \in \Omega_n$ such that $BR(\tau) = R(\pi)B$.

The explicit matrix form of linear transformations A and B satisfying Conditions 1 and 2 is given by Theorem 1 and Corollary 1.1 in the next section. Theorem 2 states that A satisfies Condition 1 if and only if A^- satisfies Condition 2. Theorems 3 and 4 show that Condition 1 is necessary and sufficient for A to preserve majorization and A^- to preserve exchangeability. In the case $m = n$, Corollary 1.3 shows that Conditions 1 and 2 are equivalent, and Theorems 5 and 6 imply that the nonsingular linear transformations satisfying these conditions preserve majorization exchangeability and Schur concavity.

2. BASIC THEORY

This section contains a series of theorems and corollaries that illustrate the ramifications of Conditions 1 and 2. Theorems 4, 5, 7, and 8 apply in the basic situation where all random variables have proper densities.

THEOREM 1. *A linear transformation $A: \mathfrak{R}^m \to \mathfrak{R}^n$ satisfies Condition* 1 *if and only if A can be represented in the form*

$$A = R(\sigma)\begin{bmatrix} A_1 \\ \vdots \\ A_t \end{bmatrix} \tag{1}$$

where $\sigma \in \Omega_n$, and for each $i = 1,\dots,t$, A_i is an $n_i \times m$ matrix ($\Sigma n_i = n$) satisfying the following conditions: If the first row of A_i contains the distinct elements $d_1,\dots,d_k$ with multiplicities $p_1,\dots,p_k$ (so that $\Sigma p_j = m$), then the n_i rows of A_i consist of the $n_i = m!/\prod p_j!$ distinct permutations of the first row.

Proof. Sufficiency: Suppose that $A: \mathfrak{R}^m \to \mathfrak{R}^n$ can be represented in the form (1), and let $\pi \in \Omega_m$ be an arbitrary permutation. Then

$$AR(\pi) = R(\sigma)\begin{bmatrix} A_1 R(\pi) \\ \vdots \\ A_t R(\pi) \end{bmatrix},$$

where, for each $i = 1,\dots,t$, postmultiplying A_i by $R(\pi)$ has the effect of permuting the elements in each row of A_i according to the permutation π. Because the rows of A_i contain all possible permutations of the elements of the first row, all these permutations again appear in the rows of $A_i R(\pi)$. Thus there exists a permutation $\omega_i \in \Omega_{n_i}$ of the rows of A_i such that

$$R(\omega_i)A_i = A_i R(\pi).$$

Define $\omega \in \Omega_n$ by $R(\omega) = \mathrm{diag}[R(\omega_1),\dots,R(\omega_t)]$, and let $\tau = \sigma\omega\sigma^{-1}$. Then for $A^* = [A'_1,\dots,A'_t]'$,

$$R(\tau)A = R(\sigma\omega\sigma^{-1})A = R(\sigma\omega\sigma^{-1})R(\sigma)A^* = R(\sigma)R(\omega)A^* = AR(\pi),$$

as required.

Necessity: Let $\mathbf{a}'_i$ be the ith row of A, and let M_i be the set of all distinct vectors of the form $\mathbf{a}'_i R(\pi)$ for some $\pi \in \Omega_m$. For each $j \neq i$, it is clear that either $M_i = M_j$ or M_i and M_j are disjoint. Let $M_1^*,\dots,M_q^*$ ($q \leqslant n$) denote all distinct members among the sets $M_1,\dots,M_n$. Defining r_i as the number of M_j equal to M_i^*, and n_i as the number of elements in M_i^*, it follows from

Condition 1 that $t_i = r_i/n_i$ is an integer. Let B_i be the $n_i \times m$ matrix whose rows consist of the n_i members of M_i^*, and let $A_1,\ldots, A_t$ $(t = \Sigma t_i)$ denote the sequence

$$B_1,\ldots, B_1, B_2,\ldots, B_2,\ldots, B_q,\ldots, B_q$$

where each B_i appears t_i times $(i = 1,\ldots, q)$. Then, clearly, A_i satisfies the condition of the theorem. It follows from Condition 1 that there is a permutation $\tau \in \Omega_n$ such that

$$R(\tau)A = \begin{bmatrix} A_1 \\ \vdots \\ A_t \end{bmatrix}.$$

Multiplying both sides of this equation by $R(\sigma)$, where $\sigma = \tau^{-1}$, then yields (1). ∎

The proof of the following corollary is similar to that of the theorem, and is therefore omitted.

COROLLARY 1.1. *A linear transformation $B: \mathfrak{R}^n \to \mathfrak{R}^m$ satisfies Condition 2 if and only if B can be represented in the form $B = [B_1,\ldots, B_t]R(\sigma)$ where $\sigma \in \Omega_n$, and for each $i = 1,\ldots, t$, B_i is an $m \times n_i$ matrix $(\Sigma n_i = n)$ satisfying the following conditions: If the first column of B_i contains the distinct elements $d_1,\ldots, d_k$ with multiplicities $p_1,\ldots, p_k$ (so that $\Sigma p_j = m$), then the n_i columns of B_i consist of the $n_i = m!/\Pi p_j!$ distinct permutations of the first column.*

The next two corollaries reveal the special structure of A and B when $m = n$.

COROLLARY 1.2. *A nonsingular linear transformation $A: \mathfrak{R}^n \to \mathfrak{R}^n$ satisfies Condition 1 if and only if A can be represented in the form*

$$A = aR(\sigma) + bJ_n, \tag{2}$$

where $\sigma \in \Omega_n$ and $a, b \in \mathfrak{R}$ $(a \neq 0,\ a \neq -nb)$.

Proof. Using the same notation as in Theorem 1, $n_i = n!/\Pi p_j! \leqslant n$ implies that either $p_1 = n$, which leads to the rows of A being multiples of e_n'

and A being singular with $a = 0$ in (2), or $\{p_1, p_2\} = \{1, n-1\}$, which leads to the general representation (2). ■

COROLLARY 1.3. *If $m = n$, then Conditions 1 and 2 are equivalent.*

Proof. Suppose that A satisfies Condition 1. By Corollary 1.2, A can be expressed as $A(\sigma) = aR(\sigma) + bJ$, where $a, b \in \Re$. Thus for every $\pi, \tau \in \Omega_n$, $A(\sigma)R(\tau) = A(\tau\sigma)$ and $R(\pi)A(\sigma) = A(\sigma\pi)$. In particular, for any $\pi \in \Omega_n$, setting $\tau = \sigma\pi\sigma^{-1}$ makes $A(\sigma)R(\tau) = R(\pi)A(\sigma)$, which means that A satisfies Condition 2. The proof of the converse is similar. ■

It follows that if B is a nonsingular linear transformation satisfying Condition 2 with $m = n$, then B must also have the form (2). A more general relationship between conditions 1 and 2, valid for all $m \leqslant n$, is given by Theorem 2.

THEOREM 2. *The linear transformation $A: \Re^m \to \Re^n$ satisfies Condition 1 if and only if its generalized inverse A^- satisfies Condition 2.*

Proof. Suppose that A satisfies Condition 1. Then, since $A^-A = I$, for every permutation $\pi \in \Omega_m$, $R(\pi) = A^-AR(\pi) = A^-R(\tau)A$ for some $\tau \in \Omega_n$. Thus $R(\pi)A^- = A^-R(\tau)AA^-$. The claim is that $A^-R(\tau)AA^- = A^-R(\tau)$.

Let N be the null space of A^-, and M be the range of A, so that $\Re^n = N \oplus M$ (see, for example, Theorem 6.11 of Graybill, 1983). For any $\mathbf{x} \in \Re^n$, let $\mathbf{x} = \mathbf{x}_N + \mathbf{x}_M$ be its unique decomposition as a sum of vectors from N and M, respectively. Note that $\mathbf{x}_M = AA^-\mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ onto M.

For any $\mathbf{x} \in \Re^n$,

$$A^-R(\tau)AA^-\mathbf{x} = A^-R(\tau)\mathbf{x}_M$$

and

$$A^-R(\tau)\mathbf{x} = A^-R(\tau)(\mathbf{x}_N + \mathbf{x}_M) = A^-R(\tau)\mathbf{x}_M + A^-R(\tau)\mathbf{x}_N.$$

Thus it remains to show that

$$A^-R(\tau)\mathbf{x}_N = 0. \tag{3}$$

Let $\Psi = \{\tau \in \Omega_n : R(\tau)A = AR(\pi) \text{ for some } \pi \in \Omega_m\}$. Then Ψ is a subgroup of Ω_n, since $\tau \in \Psi$ implies $R(\tau^k)A = AR(\pi^k)$ for each $k = 1, 2, \ldots,$

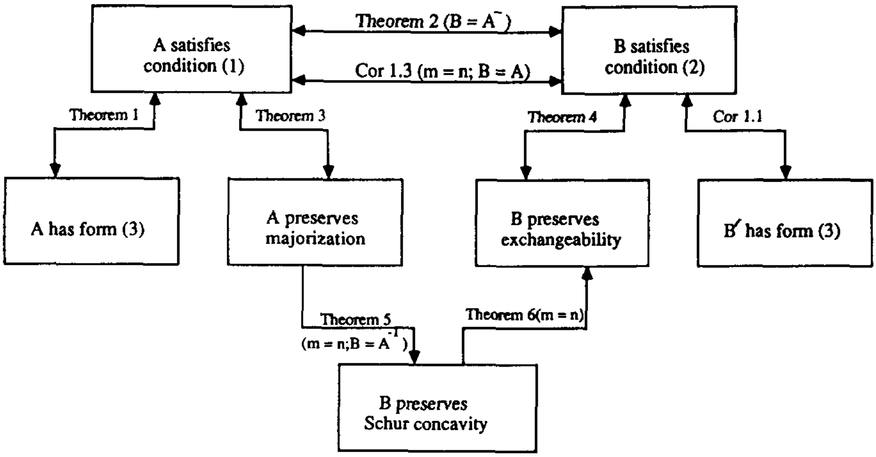

FIG. 1. Schematic of basic results.

and hence that $\tau^{-1} \in \Psi$; and, if τ_1 and $\tau_2 \in \Psi$ correspond to π_1 and $\pi_2 \in \Omega_m$, then $R(\tau_1\tau_2)A = R(\tau_1)R(\tau_2)A = R(\tau_1)AR(\pi_2) = AR(\pi_1)R(\pi_2) = AR(\pi_1\pi_2)$.

Both N and M are invariant under Ψ, as is now demonstrated. Let $x_M \in M$ and $\tau \in \Psi$. Then $R(\tau)x_M = R(\tau)Ay = AR(\pi)y$ for some $y \in \Re^m$ and $\pi \in \Omega_m$, implying that $R(\tau)x_M \in M$. Thus M is invariant under Ψ. For $x_N \in N$, $x_M \in M$, and $\tau \in \Psi$, we have $[R(\tau)x_N]'x_M = x_N'[R(\tau)]'x_M = x_N'R(\tau^{-1})x_M = 0$, since $\tau^{-1} \in \Psi$, M is invariant under Ψ, and N is orthogonal to M. Thus $R(\tau)x_N \in N$, since N is the orthogonal complement of M; hence N is invariant under Ψ.

Since the τ in Equation (3) is in Ψ, it follows that $R(\tau)x_N \in N$, and hence that Equation (3) is satisfied.

The proof of the converse is similar, and is thus omitted. ■

The relationship between pairs of linear transformations satisfying Conditions 1 and 2 and those preserving majorization, Schur concavity, and exchangeability will now be developed in a series of theorems, according to Figure 1.

THEOREM 3. *A linear transformation* $A: \Re^m \to \Re^n$ *satisfies Condition* 1 *if and only if it preserves majorization.*

Proof. Sufficiency: Suppose that A preserves majorization. Since for any $\pi \in \Omega_m$ and $y \in \Re^m$, y majorizes and is majorized by $R(\pi)y$, it follows that

$A\mathbf{y}$ majorizes and is majorized by $AR(\pi)\mathbf{y}$. Such a reciprocal relationship occurs if and only if each row of $AR(\pi)\mathbf{y}$ appears as a row of $A\mathbf{y}$. Therefore there is a permutation $\tau \in \Omega_n$ such that $R(\tau)A\mathbf{y} = AR(\pi)\mathbf{y}$. Thus $R(\tau)A = AR(\pi)$, and Condition 1 holds.

Necessity: If $\mathbf{y}$ and $\mathbf{z} \in \mathfrak{R}^m$ are ordered by majorization as $\mathbf{y} > \mathbf{z}$, then there is some $D \in \mathscr{D}_m$, the set of all $m \times m$ doubly stochastic matrices, such that $\mathbf{z} = D\mathbf{y}$. By Birkoff's theorem (Marshall and Olkin, 1979, Chapter 2) any $D \in \mathscr{D}_m$ can be represented as a convex combination of permutation matrices $D = \Sigma w_i R(\pi_i)$, where $\pi_i \in \Omega_m$, $w_i \geqslant 0$, and $\Sigma w_i = 1$. Thus $\mathbf{y} > \mathbf{z}$ implies that

$$\mathbf{z} = \sum w_i R(\pi_i)\mathbf{y}.$$

If a linear transformation $A: \mathfrak{R}^m \to \mathfrak{R}^n$ satisfies Condition 1, then

$$\begin{aligned} A\mathbf{z} &= \sum w_i AR(\pi_i)\mathbf{y} \\ &= \sum w_i R(\tau_i)A\mathbf{y} \end{aligned}$$

for some $\tau_i \in \Omega_n$; but because $\Sigma w_i R(\tau_i) \in \mathscr{D}_n$, it follows that $A\mathbf{y} > A\mathbf{z}$. ■

THEOREM 4. *A linear transformation* $B: \mathfrak{R}^n \to \mathfrak{R}^m$ *preserves exchangeability if and only if it satisfies Condition* 2.

Proof. Sufficiency: Suppose that $\mathbf{x} \in \mathfrak{R}^n$ has exchangeable components and that $\mathbf{y} = B\mathbf{x}$, where B satisfies Condition 2. Then for all $\pi \in \Omega_m$, there exists a $\tau \in \Omega_n$ such that $R(\pi)\mathbf{y} = R(\pi)B\mathbf{x} = BR(\tau)\mathbf{x}$. Since $R(\tau)\mathbf{x}$ has the same distribution as $\mathbf{x}$, $R(\pi)\mathbf{y}$ must have the same distribution as $B\mathbf{x} = \mathbf{y}$; hence $\mathbf{y}$ has exchangeable components and B preserves exchangeability.

Necessity: Let $\mathbf{x}_0 \in \mathfrak{R}^n$ be a fixed vector with distinct components, and define the probability mass function f on $\mathfrak{R}^n$ by $f(\mathbf{x}) = 1/n!$ if $\mathbf{x} = R(\tau)\mathbf{x}_0$ for some $\tau \in \Omega_n$, and $f(\mathbf{x}) = 0$ otherwise. Then $\mathbf{x} \sim f$ has exchangeable components.

Suppose that $\mathbf{y} = B\mathbf{x}$ where $B: \mathfrak{R}^n \to \mathfrak{R}^m$ preserves exchangeability. The probability mass function g of $\mathbf{y}$ has support on $S = \{\mathbf{z} \in \mathfrak{R}^m : \mathbf{z} = BR(\tau_1)\mathbf{x}_0$ for some $\tau_1 \in \Omega_n\}$. By assumption, B preserves exchangeability, and therefore $R(\pi)S \equiv \{R(\pi)\mathbf{z} : \mathbf{z} \in S\} = S$ for all $\pi \in \Omega_m$. Consequently, for all $\pi \in \Omega_m$ and $\tau_1 \in \Omega_n$, there exists a $\tau_2 \in \Omega_n$ such that $R(\pi)BR(\tau_1)\mathbf{x}_0 = BR(\tau_2)\mathbf{x}_0$. Conceivably τ_2 could depend on the particular vector $\mathbf{x}_0$ chosen; however, since vectors with distinct components are dense in $\mathfrak{R}^n$, it is possible to find a spanning set of such vectors $\mathbf{x}_0$ for which the same τ_2 is chosen. It follows that $R(\pi)BR(\tau_1) = BR(\tau_2)$ and that $R(\pi)B = BR(\tau)$ for $\tau = \tau_2\tau_1^{-1}$. ■

THEOREM 5. *If the nonsingular linear transformation $A: \Re^n \to \Re^n$ preserves majorization, then A^{-1} preserves Schur concavity.*

Proof. Let f be the probability density function of $\mathbf{x}$, so that the density function g of $\mathbf{y} = A^{-1}\mathbf{x}$ is given by

$$g(\mathbf{y}) = |\det(A)|f(A\mathbf{y}).$$

If $\mathbf{y}_1 > \mathbf{y}_2$ then $A\mathbf{y}_1 > A\mathbf{y}_2$ because A preserves majorization; and if f is Schur-concave, $g(\mathbf{y}_1)/g(\mathbf{y}_2) = f(A\mathbf{y}_1)/f(A\mathbf{y}_2) \leqslant 1$, ensuring that g is Schur-concave also. Thus A^{-1} preserves Schur concavity. ■

The following lemma describes a key property about linear transformations that preserve Schur concavity.

LEMMA 1. *If a nonsingular linear transformation $B: \Re^n \to \Re^n$ preserves Schur concavity, then $\mathbf{e}_n$ is an eigenvector of B^{-1}.*

Proof. Let $\|\mathbf{x}\|^2 = \Sigma|\mathbf{x}_i|^2$ be squared Euclidean distance, and for each $\varepsilon > 0$, let $S_\varepsilon = \{\mathbf{x} \in \Re^n : \|\mathbf{x} - n^{-1}\mathbf{x}'\mathbf{e}_n\mathbf{e}_n\|^2 \leqslant \varepsilon^2\}$ be the cylinder of radius ε about the line determined by $\mathbf{e}_n$. Since S_ε is S-convex (cf. Marshall and Olkin, 1979, p. 71), it is possible to construct a Schur-concave density f_ε with support S_ε. For example, if $\mathbf{x}$ has a standard multivariate normal density, then the conditional density of $\mathbf{x}$ given that $\mathbf{x} \in S_\varepsilon$ is such a density.

Let $\mathbf{x} \sim f_\varepsilon$ and $\mathbf{y} = B\mathbf{x} \sim g_\varepsilon$. Because B is nonsingular, the support of $g_\varepsilon = B(S_\varepsilon)$ contains some point $\mathbf{z}_\varepsilon$ such that $c_\varepsilon = n^{-1}\mathbf{z}_\varepsilon'\mathbf{e}_n \neq 0$. By assumption, B preserves Schur concavity, so that g_ε must be Schur-concave. Thus $g_\varepsilon(c_\varepsilon\mathbf{e}_n) \geqslant g_\varepsilon(\mathbf{z}_\varepsilon) > 0$, and $c_\varepsilon\mathbf{e}_n \in B(S_\varepsilon)$ is a nonzero scalar multiple of $\mathbf{e}_n$.

Suppose that as $\varepsilon \to 0$, some sequence $\mathbf{z}_\varepsilon$ can be chosen so that for each ε, $|c_\varepsilon| > \delta$ for some $\delta > 0$. If $\mathbf{v} = B^{-1}(\mathbf{e}_n)$ were not a scalar multiple of $\mathbf{e}_n$, then it would be possible to choose ε small enough so that $\delta\mathbf{v} \notin S_\varepsilon$, contradicting the fact that $B(c_\varepsilon\mathbf{v}) = c_\varepsilon\mathbf{e}_n \in B(S_\varepsilon)$. Thus, in this case, $\mathbf{v} = B^{-1}(\mathbf{e}_n)$ must be a scalar multiple of $\mathbf{e}_n$.

Alternatively, if the above supposition does not hold, then $c_\varepsilon \to 0$ as $\varepsilon \to 0$ no matter what choices are made for each $\mathbf{z}_\varepsilon$. In this case $\mathbf{z}_\varepsilon'\mathbf{e}_n \to 0$ as $\varepsilon \to 0$ implies that any vector $\mathbf{z}$ which is in $B(S_\varepsilon)$ for every $\varepsilon > 0$ must be orthogonal to $\mathbf{e}_n$. In particular, $\mathbf{y}_1 = B\mathbf{e}_n$ must be orthogonal to $\mathbf{e}_n$, so that $\mathbf{y}_1 > \mathbf{0}$. Let C be any compact S-convex subset of $\Re^n$ containing $\mathbf{e}_n$ and $\mathbf{0}$ and having positive Lebesgue measure. Define the Schur-concave density $f(\mathrm{x})$ to be proportional to $\exp\{|\mathrm{x}'\mathrm{e}_n|\}$ on C and zero elsewhere, and let $g(\mathrm{y})$ be the corresponding density of $\mathrm{y} = B\mathrm{x}$. Then $g(\mathrm{y}_1)/g(\mathbf{0}) = f(\mathrm{e}_n)/f(\mathbf{0}) = e^n > 1$ contradicts the requirement that g be Schur-concave, since B was assumed to preserve Schur concavity. This contradiction shows that $\mathrm{y}_1 = B\mathrm{e}_n$ cannot be orthogonal to e_n if B is to preserve Schur concavity; hence the possibility in the preceding paragraph must hold, namely that $\mathrm{v} = B^{-1}(\mathrm{e}_n)$ must be a scalar multiple of e_n, making e_n an eigenvector of B^{-1}. ■

THEOREM 6. *If a nonsingular linear transformation $B: \mathfrak{R}^n \to \mathfrak{R}^n$ preserves Schur concavity, then it also preserves exchangeability.*

Proof. Let $\mathrm{v}_i = (0,\dots,1,\dots,0)'$ be the unit vector in $\mathfrak{R}^n$ with ith component equal to 1, and let S be the convex hull of $\{\mathrm{v}_1,\dots,\mathrm{v}_n\}$. Note that S is S-convex, so that

$$f(\mathrm{x}) = \begin{cases} c\exp\left(-\sum x_j^2\right) & \text{for } \mathrm{x}\in S, \\ 0 & \text{for } \mathrm{x}\notin S \end{cases}$$

is Schur concave. Let the scalar c be chosen so that f is a proper density, let $\mathrm{x} \sim f$, and let $\mathrm{y} = B\mathrm{x} \sim g$. Since B is linear, the support $B(S)$ of g is the convex hull of $\{\mathrm{w}_1,\dots,\mathrm{w}_n\}$, where $\mathrm{w}_i = B\mathrm{v}_i$, $i = 1,\dots,n$. By assumption, B preserves Schur concavity, so that g must be Schur concave. Hence for any permutation $\tau\in\Omega_n$, $g[R(\tau)\mathrm{w}_i] = g(\mathrm{w}_i)$. Moreover, since f takes its minimal value only on $\{\mathrm{v}_1,\dots,\mathrm{v}_n\}$, g takes its minimal value only on $\{\mathrm{w}_1,\dots,\mathrm{w}_n\}$. Thus $R(\tau)\mathrm{w}_i\in\{\mathrm{w}_1,\dots,\mathrm{w}_n\}$. Furthermore, Lemma 1 implies that $\mathrm{w}_i \neq \mathrm{e}_n$, since $\mathrm{v}_i = B^{-1}\mathrm{w}_i$ is not a scalar multiple of e_n. Thus there is some $\tau\in\Omega_n$ such that $R(\tau)\mathrm{w}_i \neq \mathrm{w}_i$. It follows that $\{\mathrm{w}_1,\dots,\mathrm{w}_n\} = \{R(\tau)\mathrm{w}_1 : \tau\in\Omega_n\}$. Let π be an arbitrary permutation in Ω_n. Then for each $i = 1,\dots,n$, $BR(\pi)\mathrm{v}_i = B\mathrm{v}_{\pi^{-1}(i)} = \mathrm{w}_{\pi^{-1}(i)} = R(\tau)\mathrm{w}_i = R(\tau)B\mathrm{v}_i$, for some $\tau\in\Omega_n$. Since $\{\mathrm{v}_1,\dots,\mathrm{v}_n\}$ are a basis for $\mathfrak{R}^n$, it follows that for every $\pi\in\Omega_n$ there is a $\tau\in\Omega_n$ such that $BR(\pi) = R(\tau)B$. This means that B satisfies Condition 1. By Corollary 1.3, B also satisfies Condition 2; and by Theorem 4, B preserves exchangeability. ■

Combining our basic results produces the following theorems.

THEOREM 7. *A linear transformation $A: \mathfrak{R}^n \to \mathfrak{R}^m$ of rank m preserves majorization, and A^- preserves exchangeability if and only if A has the form* (1).

THEOREM 8. *A nonsingular linear transformation* $A: \mathfrak{R}^n \to \mathfrak{R}^n$ *preserves majorization, Schur concavity, and exchangeability if and only if* A *has the form* (2).

3. PARTIAL EXCHANGEABILITY

Many of the basic results in Section 2 have simple extensions in terms of partial exchangeability. In this section we limit our discussion of such generalizations to those results needed to prove Theorem 10 below, which is useful for determining the partial exchangeability of certain sets of estimators of interactions in a general linear model with exchangeable errors.

DEFINITION. Let Φ be a subgroup of Ω_m. A random vector $\mathbf{y} \in \mathfrak{R}^m$ has *Φ-partially exchangeable components* if $R(\pi)\mathbf{y}$ has the same distribution as $\mathbf{y}$ for each $\pi \in \Phi$. A linear transformation $B: \mathfrak{R}^n \to \mathfrak{R}^m$ *preserves Φ-partial exchangeability* if $\mathbf{y} = B\mathbf{x}$ has Φ-partially exchangeable components whenever $\mathbf{x}$ has exchangeable components.

The proof of Theorem 9 is similar to that of Theorem 6 and is thus omitted.

THEOREM 9. *For Φ a subgroup of Ω_m, the linear transformation $B: \mathfrak{R}^n \to \mathfrak{R}^m$ preserves Φ-partial exchangeability if and only if it satisfies the following condition:*

CONDITION 2′. *For every permutation $\pi \in \Phi$ there exists a permutation $\tau \in \Omega_n$ such that $BR(\tau) = R(\pi)B$.*

The next theorem generalizes a result by Dean and Wolfe (1988), which was used to determine distributional properties of linear estimators in a nonparametric linear model. The notation is as follows. For $i = 1,2$, C_i is an $m_i \times n_i$ matrix of rank $m_i < n_i$, I_i is the $n_i \times n_i$ identity matrix, and $\mathbf{e}_i' = [1, \dots, 1]$ is a vector with n_i components.

THEOREM 10. *Let $\mathbf{x}$ be a random vector with $n = n_1 n_2$ exchangeable components. If $\mathbf{y}_1 \equiv [C_1 \otimes \mathbf{e}_2']\mathbf{x}$ and $\mathbf{y}_2 \equiv [\mathbf{e}_1' \otimes C_2]\mathbf{x}$ both have exchangeable components, then $\mathbf{y} \equiv [C_1 \otimes C_2]\mathbf{x}$ has Φ-partially exchangeable components for $\Phi = \{\pi \in \Omega_m : R(\pi) = R(\pi_1) \otimes R(\pi_2)$ for some $\pi_i \in \Omega_{m_i}$, $i = 1,2$, with $m = m_1 m_2\}$.*

Proof. If $\mathbf{x}$ has $n = n_1 n_2$ exchangeable components, then $\mathbf{x}_1 = [I_1 \otimes \mathbf{e}_2']\mathbf{x}$ has n_1 exchangeable components, since $I_1 \otimes \mathbf{e}_2'$ satisfies Condition 2. Similarly, $\mathbf{x}_2 = [\mathbf{e}_1' \otimes I_2]\mathbf{x}$ has n_2 exchangeable components.

By assumption, $\mathbf{y}_1 = [C_1 \otimes \mathbf{e}_2']\mathbf{x}$ has m_1 exchangeable components, so that $[C_1 \otimes \mathbf{e}_2']\mathbf{x} = C_1[I_1 \otimes \mathbf{e}_2']\mathbf{x} = C_1 \mathbf{x}_1$ has m_1 exchangeable components. Therefore, from Theorem 4, C_1 satisfies Condition 2 with the appropriate dimensions, and for similar reasons so does C_2. Consequently, if $\pi \in \Omega_m$ is such that $R(\pi) = R(\pi_1) \otimes R(\pi_2)$ for some $\pi_i \in \Omega_{m_i}$ $(i = 1,2)$, then there exist $\tau_1 \in \Omega_{n_1}$ and $\tau_2 \in \Omega_{n_2}$ such that

$$\begin{aligned} R(\pi)[C_1 \otimes C_2] &= R(\pi_1)C_1 \otimes R(\pi_2)C_2 \\ &= C_1 R(\tau_1) \otimes C_2 R(\tau_2) \\ &= [C_1 \otimes C_2][R(\tau_1) \otimes R(\tau_2)], \end{aligned}$$

and Condition 2′ is satisfied for $\Psi_m = \{\pi \in \Omega_m : R(\pi) = R(\pi_1) \otimes R(\pi_2)\}$. ∎

COROLLARY 10.1. *If* $n = n_1 n_2 \cdots n_p$, *and* $\mathbf{y}_j = [\mathbf{e}_1' \otimes \cdots \otimes C_j \otimes \cdots \otimes \mathbf{e}_p']\mathbf{x}$ *has* m_j *exchangeable components* $(j = 1, \dots, p)$, *then* $\mathbf{y} = [C_1 \otimes \cdots \otimes C_p]\mathbf{x}$ *has* Φ*-partially exchangeable components for* $\Phi = \{\pi \in \Omega_m : R(\pi) = R(\pi_1) \otimes \cdots \otimes R(\pi_p)$ *for some* $\pi_i \in \Omega_{m_i}$, $i = 1, \dots, p\}$, *with* $m = m_1 m_2 \cdots m_p$.

EXAMPLE. For the linear model $\mathbf{x} = \boldsymbol{\mu} + \mathbf{e}$, where $\boldsymbol{\mu}$ is a vector of unknown constants and $\mathbf{e} \in \Re^n$ has exchangeable components, consider a factorial experiment with three factors having $n_1 = 3$, $n_2 = 4$, $n_3 = 5$ levels, respectively, and one observation on each combination of factor levels. Suppose that the levels of the first factor are qualitative and that the effect of the first level is to be compared with the effects of the other two levels of that factor. The contrasts of interest may then be written as $B_1\boldsymbol{\mu} = [C_1 \otimes \mathbf{e}_2' \otimes \mathbf{e}_3']\boldsymbol{\mu}$, where

$$C_1 = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{bmatrix}.$$

Suppose that the second factor has four levels. A set of contrasts that compares the levels of the second factor in pairs is $B_2\boldsymbol{\mu} = [\mathbf{e}_1' \otimes C_2 \otimes \mathbf{e}_3']\boldsymbol{\mu}$, where

$$C_2 = \begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.$$

As a final variation, suppose that for the third factor, it is of interest to compare the effect of the first level with the average of the other four levels, and also to compare the fifth level with the average of the other four levels. The contrasts of interest are then $B_3\mu = [e_1' \otimes e_2' \otimes C_3]\mu$, where

$$C_3 = \begin{bmatrix} 4 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 4 \end{bmatrix}.$$

As is well known, the least-squares estimators of the $B_i\mu$ are $B_i\mathbf{x}$ ($i = 1,2,3$). Under the null hypothesis that $\mathbf{x}$ has exchangeable components, Theorem 4 implies that each $B_i\mathbf{x}$ has exchangeable components.

The interaction contrasts $[e_1' \otimes C_2 \otimes C_3]\mu$, $[C_1 \otimes C_2 \otimes e_3']\mu$, $[C_1 \otimes e_2' \otimes C_3]\mu$, and $[C_1 \otimes C_2 \otimes C_3]\mu$ correspond to the chosen sets $\{B_i\}$ of main contrasts and are estimated by replacing μ with $\mathbf{x}$. Under the null hypothesis that $\mathbf{x}$ has exchangeable components, Corollary 10.1 implies that each $B_i\mathbf{x}$ has partially exchangeable components.

4. SINGULAR DENSITIES

Suppose that $\mathbf{x} \in \mathfrak{R}^n$ has a singular density concentrated on the hyperplane $B(\mathfrak{R}^n)$, where B is an $m \times n$ matrix with $m < n$. In the case where $\mathbf{x}$ has exchangeable components and $B(\mathbf{x})$ has a Schur-concave density, it is possible to obtain a stochastic ordering of the same form as that given by the main theorem in Marshall and Olkin (1974), reproduced as Theorem 11 below. Their theorem applies to the special case where $\mathbf{x}$ has a proper density function with respect to Lebesgue measure on $\mathfrak{R}^n$. Theorem 12 generalizes Theorem 11 to include cases where $\mathbf{x}$ has a singular density, and Theorem 13 provides alternative conditions that guarantee the same conclusions as those of Theorem 11.

THEOREM 11 (Marshall and Olkin, 1974). *If $\mathbf{x} \in \mathfrak{R}^n$ has a Schur-concave density, μ is a nonrandom vector in $\mathfrak{R}^n$, and S is an S-convex subset of $\mathfrak{R}^n$, then $h(\mu) \equiv P\{\mathbf{x} \in S + \mu\}$ is a Schur-concave function of μ.*

The following lemma is needed to prove two different extensions of Theorem 11 that are valid in the case where $\mathbf{x} \in \mathfrak{R}^n$ does not have a Schur-concave density.

LEMMA 2. *If S is an S-convex subset of $\mathfrak{R}^n$ and $B: \mathfrak{R}^n \to \mathfrak{R}^m$ satisfies Condition 2, then $L = B(S)$ is an S-convex subset of $\mathfrak{R}^m$.*

Proof. First note that any subset T of $\Re^m$ is S-convex if and only if for every $y \in T$ the convex hull

$$C(y) = \left\{ \sum_{\pi \in \Omega_m} a_\pi R(\pi) y : a_\pi \geqslant 0 \text{ and } \sum_{\pi \in \Omega_m} a_\pi = 1 \right\}$$

of the orbit of y is contained in T (see Marshall and Olkin, 1979, p. 8).

Let $y \in L = B(S)$, so that $y = Bx_0$ for some $x_0 \in S$. Any element in $C(y)$ has the form

$$\begin{aligned} \sum a_\pi R(\pi) y &= \sum a_\pi R(\pi) B x_0 \\ &= \sum a_\pi B R(\tau_\pi) x_0 \qquad \text{(by Condition 2)} \\ &= B\left[\sum a_\pi R(\tau_\pi) x_0\right]; \end{aligned}$$

but $\sum a_\pi R(\tau_\pi) x_0 \in C(x_0)$ and so must also be in S because S is S-convex. Thus, $C(y)$ is a subset of L, and L is S-convex. ■

Now suppose that x has a singular density that is Schur-concave on an m-dimensional hyperplane spanned by vectors of the form $y = Bx$, where the linear transformation B satisfies some suitable condition such as Condition 2. Thwarting a simple extension of Theorem 11 in this case is the fact that majorization may be completely reversed in mapping from $\Re^n$ to the hyperplane. For example, with $n = 4$ and $m = 2$ and $B = [I_2 \ 0]$, a 2×4 matrix, $[1,1,4,0]'$ strictly majorizes $[2,0,2,2]'$, but $B[1,1,4,0]' = [1,1]'$ is strictly majorized by $B[2,0,2,2]' = [2,0]'$. Thus $h(\mu) \equiv P\{x \in S + \mu\}$ may not be Schur-concave in μ, even when $h_B(B\mu) \equiv P\{Bx \in B(S) + B\mu\}$ is Schur-concave in $B\mu$.

One possible resolution to this problem is to weaken the notion of majorization. Toward this end, we invoke the notion of group majorization used by Mudholkar (1966).

Definition. Let Ψ be a subgroup of Ω_n. A vector $w \in \Re^n$ is said to *Ψ-majorize* another vector $x \in \Re^n$ if x lies in the convex hull of $\{R(\psi)w : \psi \in \Psi\}$.

Note that in the case where $\Psi = \Omega_n$, Ψ-majorization becomes ordinary majorization, as follows from a simple application of Birkoff's theorem (see Marshall and Olkin, 1979, p. 8).

LEMMA 3. *Let $B: \mathfrak{R}^n \to \mathfrak{R}^m$ be a linear transformation satisfying Condition 2, and let $\Psi = \{\psi \in \Omega_n : BR(\psi) = R(\pi)B$ for some $\pi \in \Omega_m\}$. If $\mathbf{x}$ is Ψ-majorized by $\mathbf{w}$ in $\mathfrak{R}^n$, then $B\mathbf{x} < B\mathbf{w}$ in $\mathfrak{R}^m$.*

Proof. Since $\mathbf{x}$ is Ψ-majorized by $\mathbf{w}$, it may be written as $\mathbf{x} = \sum_{\psi \in \Psi} a_\psi R(\psi)\mathbf{w}$, for some nonnegative constants a_ψ summing to 1. Thus

$$B\mathbf{x} = \sum_{\psi \in \Psi} a_\psi BR(\psi)\mathbf{w}$$

$$= \sum_{\psi \in \Psi} a_\psi R[\pi(\psi)] B\mathbf{w},$$

where $\pi(\psi)$ is any $\pi \in \Omega_m$ for which $BR(\psi) = R(\pi)B$ and is guaranteed to exist by the definition of Ψ. Thus $B\mathbf{x}$ is in the convex hull of $\{R(\pi)B\mathbf{w} : \pi \in \Omega_m\}$, and $B\mathbf{x} < B\mathbf{w}$. ■

THEOREM 12. *Suppose that $\mathbf{y} \in \mathfrak{R}^m$ has Schur-concave density, that $A: \mathfrak{R}^m \to \mathfrak{R}^n$ satisfies Condition 1, and that $\mathbf{x} = A(\mathbf{y})$. If S is an S-convex subset of $\mathfrak{R}^n$ and $\mu \in \mathfrak{R}^n$, then*

$$h(\mu) \equiv P\{\mathbf{x} \in \mathrm{S} + \mu\} \geqslant h(\mu_0)$$

whenever μ is Ψ-majorized by μ_0 for Ψ defined as in Lemma 3.

Proof. Let $L = A^-(S)$, $\lambda = A^-(\mu)$, and $\lambda_0 = A^-(\mu_0)$. Then $h(\mu) = P\{\mathbf{x} \in \mathrm{S} + \mu\} = P\{\mathbf{y} \in L + \lambda\} \geqslant P\{\mathbf{y} \in L + \lambda_0\}$ by Theorem 11, since L is S-convex by Lemma 2 and $\lambda_0 > \lambda$ by Lemma 3; but $P\{\mathbf{y} \in L + \lambda_0\} = P\{\mathbf{x} \in \mathrm{S} + \mu_0\} = h(\mu_0)$. ■

Note that in the case $m = n$, $\Psi = \Omega_n$, and Theorem 12 reduces to Theorem 11. In the case where $\mathbf{x}$ maintains its exchangeability even when its density is singular, a stronger conclusion is given by Theorem 13.

THEOREM 13. *Suppose that $B: \mathfrak{R}^n \to \mathfrak{R}^m$ has the form $B = [B_1, \dots, B_t]R(\sigma)$ given by Corollary 1.1, B_1 is an $m \times m$ matrix, and $B_2, \dots$ Bt are $m \times 1$. Suppose also that $\mathbf{x} \in \mathfrak{R}^n$ has exchangeable components and that for any measurable subset S of $\mathfrak{R}^n$, $P\{\mathbf{x} \in \mathrm{S}\} = P\{\mathbf{y} \in B(\mathrm{S})\}$, where $\mathbf{y} = B\mathbf{x} \in \mathfrak{R}^m$ has a Schur-concave density. If S is an S-convex subset of $\mathfrak{R}^n$ and $\mu \in \mathfrak{R}^n$, then $h(\mu) \equiv P\{\mathbf{x} \in \mathrm{S} + \mu\}$ is a Schur-concave function of μ.*

Proof. We have to show that if $\mu < \nu$ then $h(\mu) \geqslant h(\nu)$. By Muirhead's theorem (Marshall and Olkin, 1979, p. 21), μ can be derived from ν by successive applications of a finite number of T-transforms of the form $T = \lambda I + (1-\lambda)Q$, where $0 \leqslant \lambda < 1$ and Q is a permutation matrix that interchanges just two coordinates and leaves the others fixed. Thus if $h(\mu) \geqslant h(\nu)$ whenever $\mu = T\nu$ for a single arbitrary T-transform, then $h(\mu) \geqslant h(\nu)$ whenever $\mu < \nu$.

Thus suppose that $\mu = T\nu$, so that μ and ν differ in exactly two coordinates determined by Q. Since $\mu < \nu$ implies that $R(\tau)\mu < R(\tau)\nu$ for any permutation $\tau \in \Omega_n$, τ may be chosen so that $\mu^* \equiv R(\sigma^{-1})R(\tau)\mu$ and $\nu^* \equiv R(\sigma^{-1})R(\tau)\nu$ differ only in the first two components, where σ is defined by $B = [B_1, \dots, B_t]R(\sigma)$. Partitioning $\mu^{*\prime} = [\mu_1', \dots, \mu_t']$ and $\nu^{*\prime} = [\nu_1', \dots, \nu_t']$ according to $[B_1, \dots, B_t]$ shows that $B_1\mu_1 < B_1\nu_1$ because B_1 preserves majorization, and that $B_i\mu_i = B_i\nu_i$ is a scalar multiple of e_m for $i = 2, \dots, t$. It follows that $B\mu^* = \sum B_i\mu_i < \sum B_i\nu_i = B\nu^*$.

If S is S-convex, then $h(\mu) = h(\mu^*) = P\{x \in S + \mu^*\}$, because x has exchangeable components and S is permutation-invariant. Thus $h(\mu) = P\{y \in L + B\mu^*\}$, where $L = B(S)$ is S-convex by Lemma 2. Applying Theorem 11, $P\{y \in L + B\mu^*\} \geqslant P\{y \in L + B\nu^*\}$, since y has a Schur-concave density, L is S-convex, and $B\mu^* < B\nu^*$. Thus $h(\mu) \geqslant h(\nu)$. ■

Under certain dimensional restrictions, the linear transformation B in Theorem 13 is guaranteed to have a block B_1 in the required form.

COROLLARY 13.1. *Suppose that B satisfies Condition 2 with $m = n-1$. Then $B = [B_1\ \ B_2]R(\sigma)$, where $\sigma \in \Omega_n$, B_1 has the form (2), and $B_2 = ce_m$ for some $c \in \Re$; and Theorem 13 applies.*

Kimura and Kakiuchi (1989) also used Muirhead's theorem to prove a special case of Theorem 13 where $B = [I_m\ \ 0_m]$ and $y = Bx$. Their purpose in developing such a theorem was to provide a stochastic inequality for certain test statistics of robust null hypotheses. The generalization provided by Theorem 13 generalizes the class of statistics known to satisfy the stochastic inequality and facilitates the verification that such test statistics satisfy the required distributional assumptions. The following example illustrates the basic idea.

EXAMPLE. Let $w \in \Re^n$ have exchangeable components, so that $x = [I_n - n^{-1}J_n]w$ also has exchangeable components and is singular, since $e_n'x = 0$. Kimura and Kakiuchi (1989) worked with statistics based on x. To check the Schur concavity of $h(\mu)$, which is needed for the distributional properties of statistics based on x, it is sufficient to check the Schur concavity of the density of $y = Bx$ for any $(n-1) \times n$ matrix B that satisfies Condition 2. Note that the Schur concavity of the density of y is easily verified if w has an n-dimensional multivariate normal density.

5. CONCLUSIONS

In this paper, we have characterized the linear transformations that preserve majorization, Schur concavity, and exchangeability. In particular, when $m = n$, we have shown that these transformations have the form (2).

For $m < n$, we have shown that any linear transformation A of the form (1) preserves majorization, and its generalized inverse $B = A^-$ preserves exchangeability. Similarly, if B' has form (1), then B preserves exchangeability, and its generalized inverse $A = B^-$ preserves majorization.

Counterparts to each of the basic theorems in Section 2 may be developed for Ψ-majorization, Ψ-partial exchangeability, and Ψ-concavity, where a function is defined to be Ψ-concave if it reverses the preordering of Ψ-majorization. The theory is quite straightforward, but its presentation would needlessly complicate the basic theory in Section 2.

We have also given two examples of the use of the basic theory developed in this paper. The first example showed that the partial exchangeability of certain contrast estimators for interaction effects in a factorial experiment can be deduced from the exchangeability of the corresponding estimators for the main effects. The second example indicated how the extension of the Marshall-Olkin theorem may be used to obtain properties of certain test statistics.

Open problems left for future research include a characterization of linear transformations $B: \Re^n \to \Re^m$ with $m < n$ that preserve Schur concavity, a complete characterization of the linear transformations that preserve partial exchangeability for a given subgroup, other applications of Theorem 13, and an investigation of a similar theory for nonlinear transformation.

We thank D. E. Critchlow, M. L. Eaton, and an anonymous referee for substantive comments that helped us improve the original draft of this paper.

REFERENCES

Anderson, T. W. 1955. The integral of a symmetric unimodal function over a symmetric convex set and some probability inequality, *Proc. Amer. Math. Soc.* 6:170–176.

Dean, A. M. and Wolfe, D. A. 1988. A Note on Exchangeability of Random Variables, Technical Report No. 395, Dept. of Statistics, Ohio State Univ.

Graybill, F. A. 1983. *Matrices with Applications in Statistics*, 2nd ed., Wadsworth, New York.

Kimura, M. and Kakiuchi, I. 1989. A majorization inequality for distributions on hyperplanes and its application to tests for outliers, *J. Statist. Plann. Inference*, to appear.

Marshall, A. and Olkin, I. 1974. Majorization in multivariate distributions, *Ann. Statist.* **2**:1189–1200.

Marshall, A. and Olkin, I. 1979. *Matrix Inequalities: The Theory of Majorization and Its Applications*, Academic, New York.

Mudholkar, G. S. 1966. The integral of an invariant unimodal function over an invariant convex set—an inequality and applications, *Proc. Amer. Math. Soc.* **17**:1327–1333.

Schur, I. 1928. Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie, *Sitzungsber. Berl. Math. Ges.* **22**:9–20.

Received 4 October 1988; final manuscript accepted 19 May 1989